Final Project Paper

College of Professional Studies,

Northeastern University

EAI6060 – Healthcare Information Processing

Professor – Philip Ballentine

02/13/2025

Title

Cardiovascular Risk Assessment in a Synthetic Patient Cohort: A Data-Driven Approach to Healthcare Improvement

Group Members

Harsh Anadkat

Abstract

This study uses a dataset of 500 individuals to examine risk factors for cardiovascular disease (CVD). Age, smoking habits, race, gender, BMI, cholesterol levels, systolic blood pressure, treatment status, and hypertension are among the characteristics included in the report. According to our data, individuals between the ages of 55 and 90 have increased rates of hypertension and possible CVD, making age a major predictor of CVD risk. The gender division in the dataset begins at row 192621 and contains both male and female patients. Additionally, racial disparities in risk variables are noted, underscoring the necessity of customized risk assessment methodologies. The purpose of this study is to offer guidance for focused CVD management and preventive tactics in clinical settings.

Introduction

Globally, cardiovascular diseases continue to be the primary cause of morbidity and death, presenting serious difficulties for healthcare and public health systems. In order to uncover important risk variables for CVD in a primary care setting, this study looks at a varied patient population. Our dataset consists of 500 patient records that include clinical measures, lifestyle variables, and demographic data. The patients, who represent a variety of racial origins, including White, Asian/Pacific Islander, Black/African American, and American Indian, range in age from newborns to the elderly. We want to identify trends that might guide more efficient, individualized approaches to CVD prevention and treatment by examining the interactions between variables including age, hypertension, BMI, cholesterol, and smoking status. Given the complexity of CVD risk and the requirement for individualized therapies in a range of patient demographics, this finding is very pertinent.

Main Body

Problem

In this study, an urban health system that serves about 500,000 people looks at the inadequate identification and management of cardiovascular disease risk in primary care settings. Even with improvements in CVD prevention, many people who are at risk go undiagnosed or receive insufficient care. The investigation specifically focuses on patients in a general practice and identifies important risk variables for adult CVD. This lessens the strain on healthcare systems by enabling early illness management (Benjamin et al., 2019).

Personas

Patients of different ages and races, primary care doctors, nurses, medical assistants, data analysts, and managers of the health system are among the many healthcare stakeholders involved in the study. Every persona adds to a thorough strategy for managing the risk of cardiovascular disease. Implementing successful data-driven solutions requires an understanding of their unique demands, difficulties, and viewpoints.

- John Smith (Male Patient): A 45-year-old White male who is overweight, a smoker, and has no known history of hypertension or diabetes. He needs clear guidance on his CVD risk and practical lifestyle modifications.
- Jane Doe (Female Patient): A 60-year-old White female with a history of hypertension. She needs support in managing her existing condition and preventing further complications.
- Dr. Alice Johnson (Physician/General Practitioner): She needs tools to efficiently identify high-risk patients, evidence-based guidelines for treatment, and actionable data insights.
- Sarah Williams (Practice Manager): She needs tools to track patient health trends, reports on practice performance, and insights for optimizing patient care.

Data

Through the use of a synthetic patient cohort, the study demonstrates complex correlations between age, diabetes, race, smoking, treatment status, hypertension, BMI, blood pressure, and cholesterol levels. A dataset of patient health data, including demographics, medical history, and lifestyle variables, is used in the study. Key variables include:

- Patient ID, age, gender, race
- Hypertension status (Y/N), Treatment status (Y/N)
- Smoking status (Y/N)
- BMI, Total Cholesterol, Systolic Blood Pressure
- Type 2 Diabetes status (Y/N)
- Cardiovascular disease status (Y/N)

Age is a significant predictor of cardiovascular risk, since patients between the ages of 55 and 90 had higher rates of hypertension and likely cardiovascular disease. Analysis must take into consideration the gender gap in the dataset, which begins at row 192621, where the majority of the patients are female. The need for customized risk assessment methods is further highlighted by the fact that risk factors vary by race.

Information and Knowledge Derived

- Age as a Risk Factor: Individuals aged 55 to 90 years showed increased risks of hypertension and possible cardiovascular disease. The prevalence of CVD and hypertension may rise with age, requiring older people to undergo more thorough screening and treatment (Benjamin et al., 2019).
- Hypertension and CVD: As previously mentioned, hypertension seems to be significantly correlated with the incidence of CVD.
- Gender-Based Differences: The dataset indicates that these factors clearly differ according to the patient's gender. These instances seem to occur more frequently in women over 55 than in males.
- Additional Risk Factors: Individuals with smoking and high BMIs were shown to have a higher incidence of CVD than nonsmokers.

According to the findings, smoking, age, BMI, and hypertension all have a significant impact on the risk of developing CVD.

Recommendations

Our data suggests that age-specific screening programs be put in place, especially for people over 55, who had higher rates of hypertension and possible CVD. Elderly individuals' hypertension should receive particular attention from healthcare professionals since it seems to be a major risk factor. Given that our data showed differences in risk variables among ethnic groups, we also recommend creating tailored treatments for various racial groups. Prioritizing lifestyle modification programs that emphasize weight management and quitting smoking is important, particularly for patients who smoke or have a high BMI because these characteristics have been linked to an increased risk of CVD. Frequent blood pressure and cholesterol checks are essential, and patients with persistently high results should get more aggressive treatment. To improve clinical decision-making and enable real-time risk assessment, we conclude by recommending the integration of these results into electronic health records. To gradually hone and enhance these tactics, ongoing data analysis and outcome monitoring should be used.

Conclusions

The intricacy of CVD risk factors and the significance of individualized methods to therapy and prevention are highlighted by this study. Age is a significant risk factor for CVD, according to our research of 500 patient records, and older patients need more thorough screening and care. The dataset's found gender inequalities, especially after row 192621, point to the necessity of gender-specific CVD preventive methods. The need of culturally appropriate therapies is underscored by racial differences in risk factors. Key areas for intervention were identified as modifiable risk factors, including smoking, body mass index, and hypertension. Healthcare professionals might possibly improve patient outcomes and lessen the overall burden of cardiovascular disease in a variety of groups by utilizing these insights to create more focused and efficient methods for CVD prevention and treatment. Validating these results in bigger cohorts and assessing the long-term effects of tailored therapies based on these risk profiles should be the main goals of future studies.

References

Retrieved from : https://figshare.com/articles/dataset/CVD_Risk_Prediction_Synthetic_Dataset/5439991?file=9399457

Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., … & American Heart Association Statistics Committee and Stroke Statistics Subcommittee. (2019). Heart disease and stroke statistics—2019 update: a report from the American Heart Association. *Circulation*, *139*(10), e56-e528.

Centers for Disease Control and Prevention (CDC). (n.d.). *Heart Disease Facts*. Retrieved from $$Insert CDC URL Here]

National Heart, Lung, and Blood Institute (NHLBI). (n.d.). *Cardiovascular Diseases*.

Laderas, Ted; Dorr, David; Vasilevsky, Nicole; McWeeney, Shannon; Haendel, Melissa; Pederson, Bjorn (2017). CVD Risk Prediction Synthetic Dataset. figshare. Dataset. https://doi.org/10.6084/m9.figshare.5439991.v1

American Heart Association. (2023). Heart disease and stroke statistics—2023 update. Circulation, 147(8), e93-e621.